DOCUMENT RESUME

ED 178 610 TM 009 890

AUTHOR Kippel, Gary M.; And Others
TITLE Improving Schools Through Research: The New York City Public School Accountability Program.
PUB DATE Apr 79
NOTE 34p.; Paper presented at the annual meeting of the American Educational Research Association (63rd, San Francisco, CA, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Bases; *Educational Accountability; Educational Assessment; *Educational Improvement; Educational Trends; Elementary Education; *Improvement Programs; Minimum Competency Testing; *Program Development; Public Schools; *Reading Achievement; Reading Tests; Success Factors; Testing Programs
IDENTIFIERS New York (New York); Test Score Decline

ABSTRACT

A research project was developed for New York City public schools to improve student achievement based on analysis of school performance and subsequent development of corrective action evaluated within a framework of accountability. The program was composed of: (1) participation in the decision making process by various special interest groups; (2) analysis of standardized achievement test data; (3) identification of exceptional schools; (4) diagnosis of school variables and conditions; and (5) development, implementation, and evaluation of corrective plans. Preliminary data were obtained from a 62-school sample. The accountability program was initiated and included a citywide computerized database of reading achievement test data for students in grades 2 through 5, and the construction of a Minimum Performance Standards Test in reading, which was directly related to the New York City public school curriculum. Exceptional schools were identified through analysis of the database and were examined to determine the educational process variables which contributed to improved student achievement. Data were reviewed by advisory committees composed of educators and parents in order to recommend improvement strategies. Accountability was defined as the responsibility to implement effective corrective action by supervisors and teachers. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED178610

Session 30.08

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Improving Schools Through Research: The New York City Public School Accountability Program

Gary M. Kippel

Joseph Reswick

Charles I. Schonhaut

Garlie A. Forehand

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Gary M. Kippel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM009 890

Running head: Improving Schools Through Research

Prepared for the Special Interest Group in Research Management, American Educational Research Association Annual Convention, San Francisco, California, April, 1979.

TABLE OF CONTENTS

Introduction 1

The Need For Improving Schools 1

New York City Public School Annual Testing Program 1

New York State Pupil Evaluation Program 2

Nationwide Indices 3

Reviews Of The Research Literature 4

Literature Surveys Conducted by the New York State Education Department 4

Additional Overviews of Related Literature 5

The New York City Public School Accountability Program 5

Design of the Accountability Program 6

Development of a Citywide Computerized Database 7

Minimum Performance Standards 8

Developing Corrective Action Plans for School Improvement. .10

The Accountability Program Field Test 13

The Sample 13

Identification of Exceptional Schools 16

Consistency of School Progress 19

Summary and Conclusion 21

Introduction

One of the most important influences on children in modern societies is the school. This educational institution is a primary agent of information transmission and socialization. In school, children learn specific disciplines such as language, mathematics, science, and the arts. In addition, they learn the speaking, reading and writing skills required for effective communication and interpersonal relations. Finally, in the school setting many of the customs, traditions and other norms of society are learned and encouraged.

The considerable responsibility invested in schools makes it especially crucial that they function with optimal effectiveness. Consequently, the following examines the need for improving schools, including an overview of related research literature. It continues by discussing New York City's public school Accountability Program which is a systematic effort to employ existing research techniques to improve schools. Subsequently, illustrative data resulting from the 62 school Accountability Program Field Test are presented.

The Need For Improving Schools

In light of the heavy responsibilities placed upon the schools, it is important to monitor school progress. Evidence of school progress, for example, is obtained by examination of trends in standardized achievement test results.

New York City Public School Annual Testing Program

Each spring the New York City Board of Education administers a citywide standardized reading achievement test to grades two through nine. Examination of Table 1 indicates that in spring, 1971 through spring, 1974, approximately 66.1, 68.0, 66.0, and 66.2 percent of students (including "Non-English Speaking") were reading below grade on forms F, H, G, and F, respectively, of the Metropolitan

Achievement Test (Colligan, et al., 1975).

Insert Table 1 about here

In spring, 1975 and spring 1976, forms C and D, respectively, of the Stanford Achievement Test were administered. Results indicated that 54.8 and 57.4 percent of students were reading below grade in 1975 and 1976, respectively (Falletta & Sheffield, 1976). In spring, 1977 the Comprehensive Test of Basic Skills was administered and results indicated that 59.9 percent of students were reading below grade level. Most recently, in spring, 1978 the California Achievement Test was administered and 57.0 percent of students were found to be reading below grade level.

When interpreting the proportions of Table 1, it should be noted that results from the abovementioned norm-referenced standardized tests must be interpreted in light of the fact that the tests are constructed so that 50 percent of the students are above and the other 50 percent are below grade level in the group used to establish the norms.[1] It is evident, however, that the majority of New York City public school students tested were below grade levels from 1971 through 1978.

New York State Pupil Evaluation Program

Each fall, the New York State Education Department administers a reading achievement test to New York City third, sixth and ninth graders as part of the Pupil Evaluation Program (University of the State of New York 1976a, 1977). The proportion of students scoring at and below the statewide reference point in reading in 1966

---

[1] It should be noted, also, that four different tests were administered on a city-wide basis through 1978. These tests were normed on different samples. Consequently, results from the different tests are not directly comparable.

and 1971 through 1978 are presented in Table 2. The statewide reference point is a score equivalent to the 23 percentile (i.e., the third stanine) obtained by all pupils in each grade tested. Unlike the distribution of scores presented in Table 1, it must be noted that the percentages released by the New York State Education Department's Pupil Evaluation Program do not include non-English speaking students.

Insert Table 2 about here

Examination of Table 2 indicates that substantial proportions of New York City pupils score at and below the statewide reference point in reading. There has been an improvement in third grade reading since 1966. In contrast, however, grade six and nine results initially declined in the periods after 1966 but have remained relatively stable during the past few years.

Nationwide Indices

On a nationwide basis, the National Assessment of Educational Progress (1976) found an increase in the average percentage of nine-year-olds answering reading items correctly when 1975 results were compared with those in 1971. However, reading performance of thirteen-and seventeen-year-olds changed little over the four year period.

Additional evidence of the need for enhanced school progress comes from both the American College Testing Program and the College Entrance Examination Board which have reported consistent declines, respectively, in American College Test batteries in English, mathematics and social studies and in Scholastic Aptitude Test verbal and mathematical subtest scores over the past decade (see, for discussion, National Institute of Education, 1976).

Reviews of the Research Literature

Substantial proportions of students reading below grade level on standardized achievement tests implies a need for enhanced school progress. The first step in the development of strategies for improving schools is a review of empirical findings to provide a firm foundation for action. Unfortunately, there is a lack of consensus regarding both the extent to which schools affect achievement and the extent to which specific variables contribute to that total influence. Some attempts to interpret the meaning and significance of the empirical literature are presented below.

Literature Surveys Conducted by The New York State Education Department

Approximately 100 studies were surveyed by the New York State Education Department's Bureau of School Programs Evaluation (University of the State of New York, 1972, 1973) in an attempt to provide information which might suggest strategies for more effective use of educational resources. Among the variables which are potentially modifiable by schools, this survey found that student performance may be related to: teacher degree status, experience, socioeconomic level, and verbal ability; number of special staff (i.e., primarily guidance counselors) in schools; and student attendance.

In a subsequent summary of studies which used three different research strategies (i.e., regression, outlier, and observation), the Bureau of School Programs Evaluation (University of the State of New York, 1976b, p. 11) concluded that "Students seem more likely to achieve well where teachers are better trained, more often certified, higher paid, and more likely to be tenured." The teachers in high performing schools, also, were warmer and more responsive to and were more supportive, had higher expectations, and more favorable perceptions of their students. It is interesting to note that this summary included student attendance rate

(once socioeconomic level was considered) and mobility as variables not associated with achievement. In addition, other variables manifesting various types of relationships with student achievement were identified.

Findings resulting from the two surveys mentioned above and two other empirical investigations (i.e., Klepak, 1974; Summers and Wolfe, 1975) were compared by New York State Assembly Education Committee Chairman Stavisky (1976). This overview indicated that different research studies have yielded varied, and at times, contradictory results.

Additional Related Literature

There have been many studies concerned with the impact of school resources on achievement. A large proportion of these investigations has been based upon, or stimulated by, "Equality of Educational Opportunity" by Coleman, et al. (1966). These include, for example, Jencks, et al. (1972), Mayeske, et al. (1972), and Mosteller and Moynihan (1972). The main consequence of these works appears to be a lack of concensus regarding both findings and analytic methods (see, for discussion of analytic methods, Pedhazur, 1975).

Two recent reviews (Averch, et al., 1972; Spady, 1973) of studies on the impact on achievement of school resources and characteristics provide a survey of findings and discuss implications. Though valuable suggestions for future research are provided, there is a lack of definitive and conclusive findings. Averch, et al. (1972, p. 154) have stated succinctly: "Research has not identified a variant of the existing system that is consistently related to students' educational outcomes."

The New York City Public School Accountability Program

In order to improve student achievement in New York City public schools, a systematic research-based plan was developed. This plan has as its goal the improve-

ment of achievement based upon an analysis of student performance and subsequent development of corrective action evaluated within a framework of accountability.

Design of the Accountability Program

In their 1969 contract, the New York City Board of Education and the United Federation of Teachers agreed to develop and implement an Accountability Program with the goal of improving student achievement. Subsequently, a broad-based Accountability Advisory Committee consisting of representatives from many civic, educational, parent, political and professional organizations was established to study and make recommendations to the Chancellor regarding accountability policy. This broad-based committee established a consensus and provided ample opportunity for participation by diverse groups with various special interests in the decision-making process.

As a result of committee recommendations, the Educational Testing Service was retained to develop a blueprint for an Accountability Program to enhance school effectiveness and thereby improve student achievement. Subsequently, "A Design for an Accountability System for the New York City School System (McDonald, et al., 1972)" was produced.

With the severe New York City fiscal crisis at the outset of the program, there were never enough funds to properly implement the program. However, the significant consensus that led to the program continued its support and active involvement.[2]

[2]Late in 1974, Community School District 17 Superintendent Charles I Schonhaut was selected as Senior Assistant to the Chancellor for Educational Performance and Accountability. In January, 1975 the Office of Educational Performance and Accountability was established as part of the Chancellor's staff and was charged with the responsibility of implementing the Accountability Program. At that time, Gary M. Kippel and Joseph Reswick were invited to join the program.

Subsequently, preliminary minimum performance standards for grade four reading were developed. Also, a longitudinal student identification system resulted in the establishment of a citywide computerized database. In addition, a field test was conducted and a method to assess school progress was developed.

The following presents an overview of the Accountability Program (see, for more detailed discussion, Schonhaut, et al., 1977, 1978). It begins with a discussion of those components which were actually accomplished: i.e., the citywide computerized database and Minimum Performance Standards. It continues with a brief discussion of plans for developing and implementing corrective action plans for school improvement. Finally, the 62 school Accountability Program Field Test is discussed.

Development of a Citywide Computerized Database

Initially, the primary focus of the Accountability Program was on norm-referenced standardized reading achievement tests. It was necessary to begin with available data and measuring instruments. As the program developed, however, the use of additional output measures, such as mathematics achievement and attendance, was anticipated. In addition, based upon availability of resources, the development of objective-referenced tests was planned.

It should be noted that standardized achievement test performance is neither the only nor the best index of student achievement and school progress. Other important student characteristics include, for example: attendance, communication skills, emotional adjustments, lateness, motivation, moral concept development, and social awareness and interpersonal relationships. A combination of measures of achievement and other growth and health indicators might provide more appropriate indices in the future.

In order to use the results of the annually administered citywide norm-referenced standardized reading achievement tests, an informational database was required.

Consequently, a computerized system was developed in order to store, retrieve and update reading achievement test data for approximately 300,000 students each year in grades two through five. The resulting database contains the test score records of both school-stable and mobile children. Initially, a unique nine-digit identification number was generated and assigned to each student in 1975. This unique number is maintained for each student throughout his or her public school career. It is recorded on both the student's cumulative test record and on computerized records.

Any student who did not have a 1975 number was assigned a new number in 1976. All grade two students, for example, are assigned a new number each year as they are tested for the first time. The primary exception in grade two is for those students who are held-over and must repeat grade two. Such students were assigned an identification number during the previous year and consequently, do not require a new number. In addition to most second-graders, students who enter the New York City public school system for the first time in later grades receive a new nine-digit identification number when they are tested for the first time.

This citywide database provides the capability of examining school progress throughout New York City. In effect, the progress of each school can be compared with that of all other schools in the city. The purpose of such comparisons is to develop ways in which schools can improve.

Minimum Performance Standards

A second source of information on school progress can be obtained by determining if a school has achieved Minimum Performance Standards. It was anticipated that corrective action plans should be applied first to schools that are below minimum performance standards.

In this sense, "standards" refer to minimally acceptable levels of student performance for instructional objectives or educational goals common to all of the schools in New York City (McDonald, et al., 1972, p. 21). Such citywide standards are established and maintained under the Chancellor's direction (McDonald, et al., 1972, p. 32).

In January, 1977 a joint $460,000 grant[3] to the New York City Board of Education and the Educational Testing Service of Princeton, New Jersey[4] funded the Minimum Performance Standards (MPS) Project. The major objective was to define standards which would provide a stimulus for schoolwide corrective action. Data collection was designed to provide a carefully documented base for the establishment of minimum standard measures in reading for grade four.

As a result of these activities, the MPS Project has developed and partially validated a measurement instrument for administration on a group basis. There are six separately reported and potentially diagnostic dimensions as follows: word recognition, word attack, word meaning, literal comprehension, reasoning comprehension and cloze comprehension.

This Minimum Performance Standards Test has been developed specifically for use and interpretation with respect to New York City public school students. It is an objective test of mastery of reading in that it has been designed to indicate which specific reading skills have been mastered. In addition, it relates directly to New York City public school curriculum. Finally, it has been carefully screened by a panel of New York City professional educators to remove possible distortion or bias and to insure its appropriateness.

---

[3] The support and encouragement of Frederic A. Mosher of the Carnegie Corporation is gratefully acknowledged.

[4] The Educational Testing Service research team was headed by Garlie A. Forehand, Director of Educational Studies, and included Lola Appel, Richard Murphy and Gita Wilder.

Developing Corrective Action Plans for School Improvement

Through the use of the citywide computerized database, it is now possible to identify schools which have achieved more or less progress than anticipated based upon previous years' test scores. In effect, schools can be assessed with respect to the extent to which their students have developed or progressed over a period of time. In the future, the combination of both the citywide database and Minimum Performance Standards can provide guidance regarding the need for, and selection of, appropriate remediation or enrichment plans.

Initially, indices of school progress are being used as preliminary indicators to classify schools for further analysis. The aim is to identify exceptional schools in which student development has been high or low, relative to the city as a whole. In the near future, comparisons can be made of what better performing schools are doing to further educational progress and, concurrently, of what is happening in poorer performing schools.

Once exceptional schools have been identified, they can be examined in greater detail in an attempt to find those educational process variables that contribute to better educational progress. Such educational process variables must be under school control. In effect, they must be variables which the school can change or about which something can be done.

Two main sources can provide information about significant process variables. One source involves the statistical analysis of existing statistical school data which can be found either in the schools themselves or at central headquarters of the New York City Board of Education. Further discussion of relevant analyses and methodology, including illustrative examples, can be found in Schonhaut, et al., (1977, 1978). The second major source involves performing case studies of exceptional schools. Such case studies can be performed by teams of experts. It is important to note that both

the statistical and case study inquiries for significant and meaningful school variables must be guided by theory and prior research.

Once those variables responsible for, or at least contributing to, school progress have been identified, they can provide the basis for corrective action in comparable schools which may be showing less progress. Corrective action refers to vigorous and effective plans and activities designed to improve school progress and student achievement. Such activities might, for example, include: instructional materials, methods of grouping students for reading, management systems and reading materials. Every school and every school district is expected to develop an action plan for improvement in response to accountability information (McDonald, et al., 1972, p. 124).

The school principal and district superintendent bear major responsibility for developing corrective action plans for the school and district, respectively. Each person who participates must know what is expected of him or her and the expectations must be realistic. To implement this, performance objectives may be developed to describe actions to be taken by teachers, administrators, counselors and paraprofessionals. The effects of implementing corrective action plans must be monitored and modifications should be instituted, as warranted.

McDonald, et al., (1972, p. 127) recommended that each school and each school district have advisory Planning and Operations working committees to insure widespread involvement in planning and in clearly defining responsibilities. Planning and Operations Committees might, for example, consist of: principal or superintendent as chairperson, officially designated teacher representative, Parent's Organization representative, and various community citizens. The precise membership might vary from school to school. In addition, the district committee would include representatives from the Community School Board.

These committees would review data provided by the central Board of Education as it pertains to their school and citywide. Each school and district would be given information as to what strategies are used by better performing schools. Each school and district might then select from these strategies what they consider most appropriate for their own corrective action plan. Eventually, each school Planning and Operations committee would develop a corrective action plan for its school and then submit it to the district and, subsequently, to the central Board of Education.

Plans would be reviewed on both levels for two purposes. First, to ascertain that it is designed to realistically improve pupil performance and second, to provide the training to implement the plan. This training must insure that each supervisor, teacher, paraprofessional, and other school personnel knows what he or she is expected to do. What each person is to do may be described as performance objectives or tasks. Obviously, required resources must be available and corrective action plans must be feasible given budgetary constraints.

Accountability, in this program, is defined in terms of the responsibility to undertake and implement effective corrective action plans. In effect, supervisors and teachers will be held accountable as to how well they carry out the corrective action plans.

In this sense, accountability should not be viewed as threatening to the staff. By gathering the data on school performance, by analyzing it, by determining what causes better performance, by developing corrective action plans, and by helping to train staff to implement such plans, the staff will have the tools to improve student performance.

The Accountability Program Field Test

In accordance with "A Design for an Accountability System for the New York City School System (McDonald, et al., 1972, pp. 167-174)," a Field Test was conducted in order to assess the feasibility of various procedures and to provide a basis for a city-wide program.

The Sample

For the Field Test, two elementary schools were selected randomly from each Community School District. The resulting 62 school sample was approximately ten percent of the elementary schools in the New York City system. This procedure was recommended by McDonald, et al., (1972, p. 168), rather than other sampling techniques, in order to provide an opportunity for each school district to participate to an equal extent at the very outset. In addition to providing useful data, an opportunity was provided for communicating the importance and usefulness of the Accountability Program to each of the school districts. It was noted that identities of students and schools in the Field Test were kept anonymous when data were analyzed and results reported.

Field Test data were collected only for students who were in grade four during the 1974-1975 academic year. The purposes of the Field Test included determination of the feasibility both of collecting retroactive data and of performing longitudinal trend analyses. Data were collected, therefore, from the 1972-1973 and 1973-1974 academic years when most of these same students were in grades two and three, respectively.

The Accountability Program is based upon the assumption that schools are responsible and accountable for effectively providing services. It is reasonable to infer that students must attend a specific school for at least a minimal amount of time for the educational program of that school to have an impact. Consequently, the focus of the Field Test is on the school-stable, rather than mobile, students (McDonald, et al.,

1972, p. 51, footnote 2). In this study, data were included only for students who were in the same school for three consecutive years.[5]

Data were collected from a three-year longitudinal sample of students who had been administered the following three norm-referenced reading tests: Stanford Achievement Test (i.e., S.A.T.)- Form 'C' Primary III level in 1975, Metropolitan Achievement Test (i.e., M.A.T.)- Form 'G' Elementary level in 1974, and the Metropolitan Achievement Test - Form 'F' Primary II level in 1973.

In order to illustrate methods for calculating school progress measures, reading comprehension raw scores were examined for 2,329 students from all 62 Accountability Program Field Test schools. Most students were in grades two, three and four in 1973, 1974 and 1975, respectively. In this sample, 2,272 progressed through these grades at the usual rate of one grade each year. The other 57 had been held-over in either grade three or four. In effect, data were retained for all school-stable students, regardless of rate of progression from grades two through four.

Descriptive statistics for 2,329 students aggregated by schools are presented in Table 3. School statistics are not based upon weighted means which take into consideration school size. It is noted that the correlation between the number of students from each school and school progress measures used in this study was found to be of zero order (i.e., r=.06). Further discussion can be found in Schonhaut, et. al., 1977, p. 74

Insert Table 3 about here

[5]It is noted, however, that an important product of the Field Test is an estimation of the extent of student mobility in Field Test schools (McDonald, et al., p. 172).

Examination of Table 3 indicates that 1975 mean raw scores aggregated by school are much higher than those for 1973 and 1974. The primary reason is that a different citywide standardized reading achievement test was administered in 1975 compared with earlier years.

It is interesting to note, also, that the school mean scores for 1973 are higher than those of 1974 even though the Metropolitan Achievement Test was administered both years. This finding results from the fact that students were administered the easier Primary II level in 1973 compared with the more difficult Elementary level in 1974.[6]

The abovementioned differences in means of the 1973, 1974 and 1975 tests are reflected, also, in differences between the ranges of student and school scores for 1973, 1974 and 1975. The standard deviations indicate least variability in 1973 (i.e., 3.270) and the greatest in 1975 (i.e., 5.227).

The correlations indicate the strongest relationship (i.e., .88), between the 1974 and 1975 tests. With the two year interval from 1973 to 1975 the correlations (i.e., .84) decreased. The weakest relationship (i.e., .81), however, was manifested between the 1973 and 1974 tests. This is interesting to note because the higher 1974-1975 correlation was based upon two different tests whereas the lower 1973-1974 correlation was based upon two different levels of the Metropolitan Achievement Test.

---

[6]Norm-referenced standardized tests, such as the Metropolitan and Stanford Achievement Tests, usually are constructed by levels appropriate for typical children in specific grades. For some children, however, the test level designed for their grade is inappropriate because test items are too easy or too difficult. In such cases, the test level designed for adjacent grades may be administered. This procedure is referred to as "out-of-level" testing and has been used at various times in the annual citywide testing program of the New York City Board of Education.

Identification of Exceptional Schools

The next phase in assessing the progress of the 62 schools was to identify exceptional schools by calculating three regression analyses on the 1973, 1974 and 1975 test scores. Regression analyses[7] are statistical techniques which can be used to identify schools that have progressed more or less than would be predicted on the basis of their past performance. Regression analyses result in regression equations which can be graphically depicted as regression lines. For example, Figure 1 illustrates the regression line resulting from the regression of 1975 reading test scores on 1974 scores. Each of the 62 dots around the regression line represents one school.

Insert Figure 1 about here

As depicted in Figure 1, the distance between the point or dot representing each school and the regression line can be thought of as an index or measure of school progress and has been referred to as the "Student Development Index (SDI)," by McDonald, et al., (1972, esp. pp. 49-55). The SDI is a measure of the educational development of students in a given school. As is suggested by the word "development," the measure involves a longitudinal comparison of student performance at a given time with student performance at an earlier time. Thus, schools are assessed with respect to the degree to which their students have progressed over a period of time. In addition, the SDI is useful in identifying exemplary schools which may serve as models for comparable schools which have not achieved as much progress.[8]

---

[7] For a more detailed review of regression methods, see Draper & Smith (1966) and Kerlinger and Pedhazur (1973).

[8] See Convey, 1977 for discussion of methods of determining school effectiveness following a regression analysis.

In Figure 1, schools represented by a point above the line have positive SDIs and have 1975 mean test scores which are higher than predicted on the basis of their 1974 mean test scores. Those schools which are relatively high above the regression line have progressed more than comparable schools. In a similar manner, those schools represented by a point below the line have negative SDIs and have 1975 mean test scores which are lower than predicted. Those schools which are relatively far below the regression line have manifested less progress than comparable schools.

The Student Development Index should not be translated as meaning "School Effectiveness Index."[9] As indicated by McDonald, et al., (1972, pp. 215-216), several factors besides the school are responsible for the influence of the formal educational process on the individual. In most cases, it is difficult or impossible to separate the school influences from those of non-school factors.[10] Therefore, it is important to note that there can be no inference that only school factors are responsible for student progress.

There are different methods for calculating SDIs (see, for discussion, Schonhaut, et al., 1977, esp. pp. 51-55). Perhaps the most important distinction between the various methods is the unit of analysis used to calculate the regression line. For example, one kind of method (see McDonald, et al., 1972, pp. 213-216) involves a regression line based upon schools. For explanatory purposes, the regression line illustrated in Figure 1 was based upon 62 schools rather than upon the 2,329 students for whom test data were aggregated.

---

[9]See, for discussion of School Effectiveness Index, Dyer (1971).

[10]For example, discussion of related methodological considerations can be found in: Armor (1972), Convey, 1973, Hanushek and Kain (1972), Pedhazur (1975), and Smith (1972).

Each analysis was based upon standardized tests which were administered only once each year. Consequently, because at least two scores were required, each analysis included scores from at least two different years. Subsequently, the analyses were designated as follows: 75/74; 75/73;and 74/73.

In each case, the most recent score or criterion variable is on the left side of the diagonal and the earlier score or predictor variable is on the right side. For example, 75/74 indicates that the 1975 test scores are predicted by 1974 test scores. In like fashion, 75/73 indicates that 1975 test scores are predicted by 1973 test scores. In this analysis, scores for each student resulted from tests which were administered approximately two years apart. In effect, scores from the intervening year (i.e., 1974) were not included. The resulting squared correlations, standard errors of estimate, regression coefficients and intercepts based upon the 62 schools are presented in Table 4.

Insert Table 4 about here

The squared correlation is an index of the percentage of variance of the latest test scores accourted for by the earlier test scores. It is a measure of regression.[11] In general, the higher the squared correlation the better the statistical model. Specifically, the higher the squared correlation the better the prediction of later scores from earlier scores. Examination of Table 4 indicates substantial squared multiple correlations ranging from .66 to .78.

---

[11]Data were examined to determine the appropriateness of the linear model. Trends beyond the linear were neither significant nor meaningful. In addition, examination of the data points and regression lines (see, for example, Figure 1) and the relatively high squared multiple correlations (see Table 4) provided evidence for the appropriateness of the linear regression model.

The standard error of estimate (SE_{est}) is the standard deviation of residuals (Kerlinger and Pedhazur, 1973, p. 66). The standard error of estimate is an index of the variation or dispersion of scores around the regression. In effect, the smaller the SE_{est} the better the prediction.

One indication that the statistical model is appropriate, and the prediction is successful, is that the SE_{est} is considerably smaller than the standard deviation of the criterion variable. As indicated by examination of Table 3, this has been achieved for the school statistics.

The regression coefficients and intercepts are used with predictor scores to comprise regression equations. These equations are used both to calculate predicted criterion scores and to draw regression lines. The regression coefficient is the slope of the line and the intercept is the point at which the line intercepts the vertical axis. For example, the regression equations from the 75/74 analysis of school data have been used to generate the regression line depicted in Figure 1. In other words, by examining the 75/74 intercept and regression coefficient presented in Table 4, we see that Figure 1 graphically depicts the equation: $y = 26.445 + 1.097x$.

Consistency of School Progress

To what extent do SDIs reflect school progress over a period of time?[12] Data were examined to assess the relationship between SDIs on the same schools calculated with one year intervening between tests (i.e., 74/73) and with two years intervening (i.e., 75/73). The resulting correlation of $r = .63$ reflects a relatively high relationship between the 74/73 and 75/73 SDIs based upon the same 62 schools.

In addition, the SDIs resulting from these two regression analyses were compared for each of the 62 schools in order to assess changes in direction or sign (i.e., positive

[12] See, for further discussion, Schonhaut, et al., 1978, p. 38.

or negative) from one year to the next. That is to say, how many schools which manifested positive measures of progress after one year had negative measures after two years? In like manner, how many schools which had negative SDIs after one year had positive SDIs after two years?

There was no change in sign (from the 74/73 to the 75/73 SDIs) for 42 of the 62 schools. Specifically, 22 schools had positive and 20 had negative SDIs for both the 74/73 and 75/73 regression analyses.

There was, however, a change in sign of SDIs for 20 of the 62 Field Test schools. Specifically, there was a positive (i.e., minus to plus) change in 13 schools and a negative (i.e., plus to minus) change in 7 schools.

Next, the consistency of SDIs for exceptional schools was examined to determine if they maintain longitudinal stability as assessed by these regression analyses. Brager (1976) has cautioned that extreme SDIs may be statistical artifacts in that they might be a product of error variance rather than due to factors of any educational significance. It may be worthwhile, therefore, to consider schools as "exceptional" only if they are consistently distant from the mean for several years (see also, Klitgaard and Hall, 1973, pp. 12-13).

Consequently, the 75/73 SDIs were examined specifically for schools which had 74/73 SDIs which were at least one standard error of estimate (i.e., one standard deviation from the mean) from the regression line. Specifically, any school which had a 74/73 SDI which exceeded plus or minus 2.49 (see Table 4) was considered as having made unusually high or low progress.

Of the six schools which had exceptionally high or positive 74/73 SDIs, two maintained exceptionally positive 75/73 SDIs. Of the nine schools which had exceptionally low 74/73 SDIs, five maintained exceptionally low 75/73 SDIs.

Overall, seven of the fifteen (i.e., 47 percent) schools with exceptional positive and negative 74/73 SDIs obtained exceptional 75/73 SDIs. In addition, of the remaining eight schools, none had 75/73 SDIs which were opposite in sign to their 74/73 SDIs. In general, these regression analyses reflect relatively consistent trends of SDIs from one year to the next.

By doing successive regression analyses each year, as illustrated by the 74/73 and 75/73 analyses, it is possible to identify some schools which manifest consistently positive or negative progress. Once such unusual schools have been identified, they can be examined systematically in greater detail in order to provide an empirical basis for corrective action.

Summary and Conclusion

There is evidence for declining reading achievement test score trends on a nationwide basis. This situation has especially ominous implications in schools and districts where many children are already reading below grade levels.

It is important for attempts to improve school effectiveness to have a systematic basis in objective and replicable research. Some valuable insights and suggestions are offered in the research literature. There is, however, a lack of general agreement among scientific investigators regarding the variables which can be expected to improve achievement.

To a large extent, this lack of consensus is understandable, considering the complexity of human functioning, limitations of research methods and procedures, and shortcomings of measuring instruments. In light of the need for improving school effectiveness, it seems crucial to apply, to whatever extent possible, existing research methods and procedures in the search for solutions.

An example was presented of one type of program which might improve school progress by: 1) providing ample opportunity for participation by diverse groups with

various special interests in the decision-making process, 2) obtaining and analyzing standardized achievement test data, 3) identifying exceptional schools, 4) diagnosing school variables and conditions, and 5) developing, implementing and evaluating corrective plans.

BIBLIOGRAPHY

Armor, D.J. School and family effects on black and white achievement: A reexamination of the USOE data. In F. Mosteller & D.P. Moynihan (Eds.), On equality of educational opportunity. New York: Vintage Books, 1972, pp. 168-229.

Averch, H.A., Carroll, S.J., Donaldson, T.S., Kiesling, H.J., & Pincus, J. How effective is schooling? A critical review and synthesis of research findings. Santa Monica, California: Rand Corporation, 1972.

Brager, G.L. Outliers and accountability: Fact or fiction? Paper presented at the meeting of the American Educational Research Association, San Francisco, California, April 1976. (ERIC Document Reproduction Service No. ED 120 230)

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfield, F.D., & York, R.L. Equality of educational opportunity. Washington, D.C.: United States Government Printing Office, 1966.

Colligan, J.M., Falletta, J., & Basso, P. Pupil reading achievement in New York City: A report of the April 1974 reading test survey grades 2 through 9 New York City public schools. Brooklyn, New York: Board of Education of the City of New York, Office of Educational Evaluation, 1975.

Convey, J.J. Some methodological considerations for school effectiveness studies. Unpublished paper, Educational Evaluation and Research Design Program, Tallahassee, Florida State University, 1973 (ERIC Document Reproduction Service No. ED 092 597)

Convey, J.J. Determining school effectiveness following a regression analysis. Journal of Educational Statistics, 1977, 2, 27-39.

Draper, N.R., & Smith, H. Applied regression analysis. New York: John Wiley, 1966.

Dyer, H.S. Toward objective criteria of professional accountability in the schools of New York City. Phi Delta Kappan, 1971, 52, 206-11.

Falletta, J., & Basso P. Pupil reading achievement in New York City: A report of the April 1975 reading test survey grades 2 through 9 New York City public schools. Brooklyn, New York: Board of Education of the City of New York, Office of Educational Evaluation, 1975.

Falletta, J.J., & Sheffield, J. Pupil reading achievement in New York City: A report of the March 1976 reading test survey grades 2 through 9 New York City public schools. Brooklyn, New York: Board of Education of the City of New York, Office of Educational Evaluation, 1976.

Falletta, J.J., & Sheffield, J. Pupil reading achievement in New York City: A report of the May 1977 reading test survey grades 2 through 9 New York City public schools. Brooklyn, New York: Board of Education of the City of New York, Office of Educational Evaluation, 1977.

Falletta, J.J., & Sheffield, J. Pupil reading achievement in New York City: A report of the March 1978 reading test survey grades 2 through 9 New York City public schools. Brooklyn, New York: Board of Education of the City of New York, Office of Educational Evaluation, 1978.

Hanushek, E.A., & Kain, J.F. On the value of Equality of Educational Opportunity as a guide to public policy. In F.Mosteller & D.P. Moynihan (Eds.), On equality of educational opportunity. New York: Vintage Books, 1972, pp.116-145.

Jencks, C., Smith, M., Acland, M.J.B., Cohen, D., Gintis, H., Heyns, B. & Michelson, S. Inequality: A reassessment of the effect of family and schooling in America. New York: Basic Books, 1972.

Kerlinger, F.N., & Pedhazur, E.J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, 1973.

Klepak, D. School factors influencing reading achievement: A case study of two inner city schools. Albany, New York: Office of Education Performance Review, State of New York, 1974.

Klitgaard, R.E., & Hall, G.R. A statistical search for unusually effective schools. Santa Monica, California: Rand Corporation, 1973.

Mayeske, G.W., Wisler, C.E., Beaton, A.E., Weinfeld, F.D., Cohen, W.M., Okada, T., Proshek, J.M., & Tabler, K.A. A study of our nation's schools. Washington, D.C.: U.S. Government Printing Office, 1972.

McDonald, F.J., Forehand, G.A., Marco, G.L., Murphy, R.T., & Quirk, T.J. A design for an accountability system for the New York City school system. Princeton, New Jersey: Educational Testing Service, 1972.

Mosteller, F., & Moynihan, D.P. (Eds.), On equality of educational opportunity. New York: Vintage Books, 1972.

National Assessment of Educational Progress. 9-year-olds improve reading skills. National Assessment of Educational Progress Newsletter, 1976, 9, 5, 1-2.

National Institute of Education. Declining test scores. Washington, D.C.: U.S.Department of Health, Education, and Welfare, 1976.

Pedhazur, E.J. Analytic methods in studies of educational effects. In F.N. Kerlinger (Ed.), Review of research in education (vol.3). Itasca, Illinois: F.E. Peacock, 1975, pp. 243-286.

Schonhaut, C.I., Kippel, G.M. & Reswick, J. Accountability program progress report. Brooklyn, New York: Board of Education, Office of Educational Performance and Accountability, October, 1977.

Schonhaut, C.I., Kippel, G.M., Koenigsberg, E., & Reswick, J. Accountability program progress report. Brooklyn, New York: Board of Education, Office of Educational Performance and Accountability, June, 1978.

Smith, M.S. Equality of Educational Opportunity: The basic findings reconsidered. In F. Mosteller & D.P. Moynihan (Eds.), On equality of educational opportunity. New York: Vintage Books, 1972, pp. 230-342.

Spady, W.G. The impact of school resources on students. F.N. Kerlinger (Ed.), Review of research in education (vol.1), Itasca, Illinois: F.E. Peacock, 1973.

Stavisky, L.P. Fact sheet prepared for public hearing on improving pupil performance in the schools; A climate for learning. Albany, New York: New York State Assembly Education Committee, 1976.

Summers, A.A., & Wolfe, B.L. Which school resources help learning? Efficiency and Equity in Philadelphia public schools. Philadelphia, Pennsylvania: Federal Reserve Bank of Philadelphia, 1975.

University of the State of New York. Variables related to student performance and resource allocation decisions at the school district level. Albany, New York: The State Education Department, Bureau of School Programs Evaluation, 1972.

University of the State of New York. What research says about improving student performance; A manual for administrators. Albany, New York: The State Education Department. Bureau of School Programs Evaluation, 1973.

University of the State of New York. The New York State Pupil Evaluation Program - 1975 summary of statewide test results in reading and mathematics. Albany, New York: State Education Department, 1976a, 1977.

University of the State of New York. Which school factors relate to learning? Summary of findings of three sets of studies. Albany, New York: The State Education Department, Bureau of School Programs Evaluation, 1976b.

Table 1. Annual New York City Standardized Reading Achievement Test Scores

Achievement Levels	Metropolitan				Stanford		Basic Skills	California
	1971	1972	1973	1974[a]	1975[b]	1976[c]	1977[d]	1978[e]
At and Above Grade	33.9	32.0	34.0	33.8	45.2	42.6	40.1	43.0
Down to One Year Below Grade	22.2	22.5	19.7	20.0	17.3	15.9	18.8	20.1
One to Two Years Below Grade	15.1	15.2	18.9	18.6	16.4	17.0	17.1	17.9
Two and More Years Below Grade	24.0	26.1	22.7	23.1	17.5	20.6	18.6	13.5
Non-English Speaking[f]	4.8	4.2	4.7	4.5	3.6	3.9	5.4	5.5

[a]Source: Colligan, et el., 1975, p.5.
[b]Source: Falletta & Basso, 1975, p.5.
[c]Source: Falletta & Sheffield, 1976, p.5.
[d]Source: Falletta & Sheffield, 1977, p.5.
[e]Source: Falletta & Sheffield, 1978, p.5.
[f]Non-English category includes students who speak little or no English or whose regional or foreign accents make it impossible for them to be readily understood.

Table 2. Proportion of New York City Pupils Scoring At and Below the Statewide Pupil Evaluation Program[a] Reference Point for Reading

	Year of Test							
Grade	1966	1971	1972	1973	1974	1975	1976	1977
Three	47	45	42	42	36	38	35	31
Six	45	50	48	51	48	49	47	46
Nine	38	--[b]	49	51	47	49	--[b]	--[b]

[a]Source: University of the State of New York, 1976a, 1977.

[b]Data not available.

Table 3. Descriptive Statistics for Schools

(N = 69)

Year	Mean	Standard Deviation	Range		Correlation	
			Minimum	Maximum	74	73
1975	56.826	5.227	39.222	65.629	.881	.839
1974	27.700	4.199	19.533	35.925		.810
1973	34.643	3.270	26.111	40.500		

Table 4. Regression Statistics for Schools

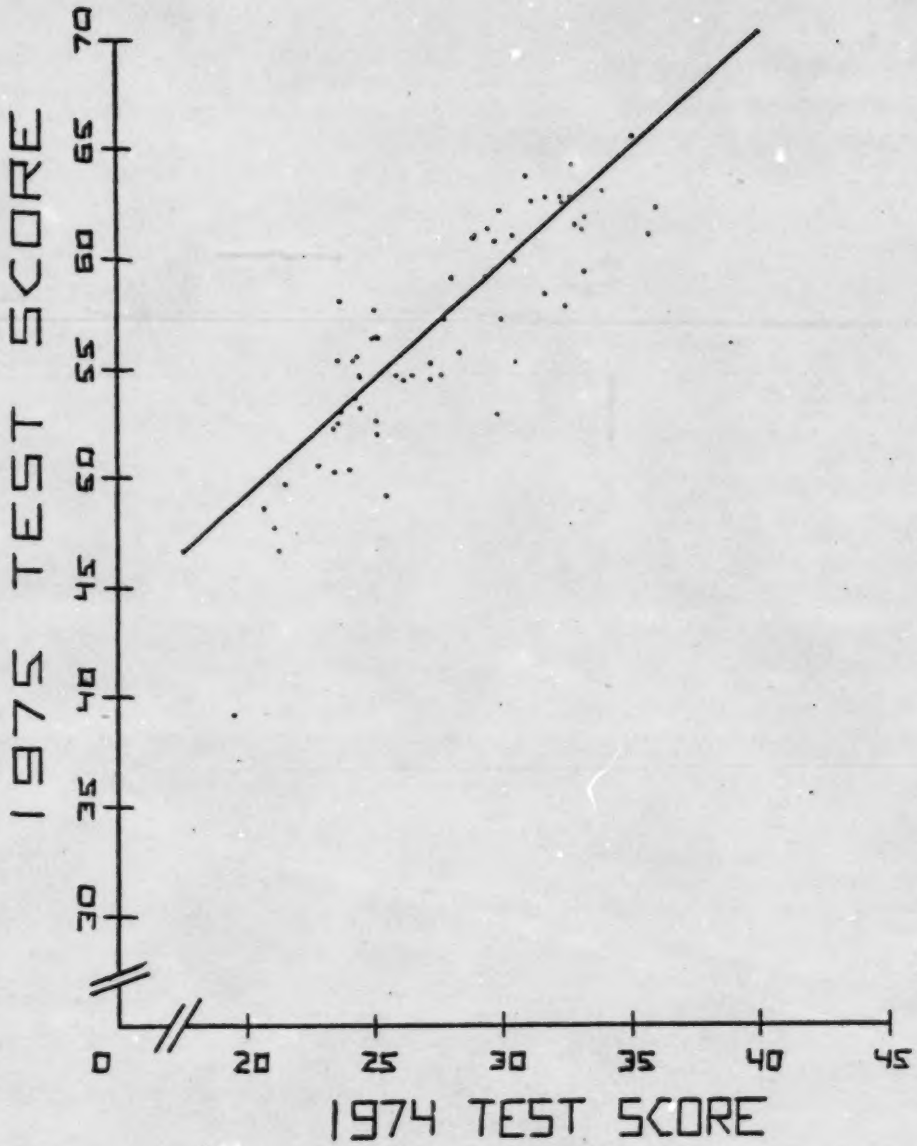
(N = 62)

Regression Analysis	Squared Correlation	Standard Error of Estimate	Regression Coefficient	Intercept
75/74	.776	2.493	1.097	26.445
75/73	.704	2.866	1.341	10.354
74/73	.656	2.485	1.040	-8.319

Figure Caption

Figure 1. Plot of 62 Accountability Program

Field Test Schools Based Upon 75/74 Regression Analysis.


1975 TEST SCORE
70
65
60
55
50
45
40
35
30
0
20
25
30
35
40
45
1974 TEST SCORE